Overreaction or Under-reaction: How Do People Respond to Wage Changes? *

Tai-Sen He†

November 30, 2013

Abstract: This paper experimentally examines individuals' work effort choices in response to wage changes. I designed a simple, real-effort experiment with four treatments. Subjects experience a wage increase in one treatment and a wage decrease in another treatment; the other two treatments are used as control treatments in comparison with the former two treatments. I find that subjects overreact to the wage increase but under-react to the wage decrease. Male subjects' average effort is higher than female subjects. However, interestingly, the wage increase and the wage decrease both have a more pronounced effect on female subjects than their male counterparts. The results have important implications for research on the wage elasticities of labor supply.

Key words: work effort, wage changes, labor supply, experiments
JEL codes: C91, D03, J31

---

* I am grateful to Louis Putterman, Pedro Dal Bó and Brian Knight for their guidance and advice. All errors are mine.

† Economics Division, Nanyang Technological University, 50 Nanyang Ave, Singapore 639798 (e-mail: TS.He@ntu.edu.sg).

I. Introduction

How do people respond to changes in wages? More precisely, how much does the labor supply respond to changes in wage rates? This classic question is central for understanding the optimal design of wage policy. Policy-makers also rely on the estimates of labor supply elasticities to evaluate the impacts of tax and transfer programs. Many studies in labor and public economics have been devoted to examining the labor supply responses to wage changes and wage-related policy changes (Blundell and MaCurdy, 1999; McClelland and Mok, 2012). Of these, nearly all of the theoretical and empirical analyses are based on standard economic models. The effect of wage change is determined by the relative sizes of the substitution effect and the income effect.

However, imagine the following scenario of two identical workers. Worker A is paid $10 per piece of work; worker B has been paid $20 per piece, but her boss suddenly cuts it down to $10. Standard economic models predict that optimal effort is determined by setting marginal cost equal to the marginal benefit defined by the piece rate (Abeler et al., 2011). Thus, these two workers should work equally hard when they are paid the same piece rate, $10. But, it is natural to doubt this, and ask: does worker B respond to wage changes as predicted?

Insights from psychology and behavioral economics provide predictions pointing in opposite directions: on the one hand, according to a stream of psychological literature on reward magnitude, worker B may feel depressed because her pay has been cut and is reluctant to work hard (Crespi effect).[1] On the other hand, recent studies in behavioral economics suggest that she may want to work harder in order to achieve the same income level as before (income target effect).[2] The Crespi effect hypothesis predicts that worker B will overreact to the wage cut and expend less effort than worker A. Conversely, the income target hypothesis would predict that worker B will under-react to the wage cut and work harder than worker A. The purpose of this paper is to examine how these

[1] Crespi effect is also known as the successive contrast effect. Please see next section for detail.

[2] Camerer et al. (1997) used data on New York cabdrivers and found evidence of negative labor supply responses to transitory wage changes. They suggested the daily targeting hypothesis based on prospect theory (Tversky and Kahneman, 1991).

behavioral factors affect an individual's work effort in the face of wage changes, both when it increases and decreases.

Laboratory experiments provide a well-controlled environment to test whether subjects over- or under-react to wage increases and decreases. I design a simple, real-effort experiment with four treatments: HH, LH, LL, and HL, shown as Figure 1. Subjects engage in a tedious keyboard-pressing task for a total of four periods. Subjects press X and Z buttons sequentially on the keyboard and their periodic earnings equal the number of XZ pairs they complete times the piece wage rate in the period. The manipulated variable is the piece wage rate: subjects are paid either at the high (1 cent per pair) or the low (0.1 cents per pair) piece wage rate. The main outcome variable of interest is the subjects' work effort level, measured by the number of XZ pairs completed in a period.

Treatment HH and LH are used to examine how subjects respond to wage increases (Figure 1 (a)). Subjects in Treatment HH are paid at the high rate in all four periods while subjects in Treatment LH are paid at the low rate in the first two periods and at the high rate in the last two periods. I compare the work effort levels between these two groups in the last two periods when subjects in both groups are paid at the same rate. The only difference between the two groups is that subjects in the treatment group (LH) have been paid at a different rate in the first two periods while the wage rate in the control group (HH) remains the same across all four periods. Similarly, Treatment LL and Treatment HL are used to examine how subjects react to wage decrease. To control for potential wealth effect, which may arise when subjects in the treatment group and in the control group are paid differently in the first two periods, subjects receive payment for one randomly drawn period only.

The main result of this paper is that individuals respond to wage changes asymmetrically: subjects overreact to the wage increase but underreact to the wage decrease. Interestingly, the wage increase and the wage decrease both have a more pronounced effect on female subjects than their male counterparts.

This paper builds on and contributes to existing literature in labor economics, analyzing labor supply decisions in response to wage changes. There have been a great number of studies devoted to estimating the wage elasticity of labor supply. However, these studies are mostly based on standard economic models, and thus the effect of wage changes is determined by the relative sizes of the substitution and income effects. This paper, on the other hand, focuses on the behavioral factors that could be triggered in the face of wage changes and uses the experimental design which allows to identify whether subjects over- or under-react to wage changes.

This paper also relates to research in behavioral economics on how labor supply is linked to various reference points, such as expectations (Abeler et al., 2011; Ericson and Fuster, 2011), goals (Heath et al., 1999) and income target (Fehr and Goette, 2007; Camerer et al., 1997). Although there has been speculation about past wage as the reference point (Feldstein and Poterba, 1984), little is known about how it acts as a reference point and how it influences labor supply decisions. Most closely related to this study is Bracha and Gneezy (2012), who experimentally examined the effect of changes in show-up fees on participation rate and found that individuals overreacted to wage changes. This paper focuses on another important aspect of labor supply, on-the-job effort provision, and find that individuals react differently depending on the direction of change.

The remainder of the paper is organized as follows. Section II presents the experimental design and theoretical predictions. Section III describes the results. I conclude in Section IV by discussing the limitations and implications of this study.

II. Experimental Design and Theoretical Predictions

2.1 Experimental Design

In my experiment, subjects engage in a keyboard-pressing task for a total of four periods, each of which lasts for four minutes. Subjects' earnings are based on the number of pairs of X and Z they complete.[3] This real-effort task is developed and used in Ariely et al. (2009) and Ariely et al. (2009). There are several advantages to this task: (1)

[3] I employ a four-period rather than a two-period design to control for a potential restart effect, which may arise in the beginning of each period. Having subjects exposed to the initial wage rate in the first two periods allows me to control for this possible restart effect. It also controls for the last-period effect: subjects have the tendency to work harder when the whole experiment is about to end.

subjects respond to financial incentives significantly; (2) it is simple and easy to communicate; (3) the task is tedious and pointless, which ensures that it entails a positive cost of effort for subjects; and (4) it requires very little cognitive ability and prior knowledge, and thus the experience effect is small.

The only experimental manipulation is the piece rate per pair of X and Z the subjects complete. Subjects are paid at either the "High" rate (1 cent per pair, for example) or the "Low" rate (0.1 cent per pair) in a single period. This large wage gap ensures that subjects behave significantly differently when paid at different piece rates (Ariely et al., 2009). The experiment is composed of two treatment groups (HL and LH) and two control groups (HH and LL) as shown in Figure 1. In the control groups, subjects are paid at a fixed piece rate, either "High" or "Low," throughout all four periods. In the treatment groups, there is an unanticipated wage change — an increase in the LH treatment and a decrease in the HL treatment — at the end of the second period.

The variable of interest is the subjects' effort level, measured by the number of X and Z pair they complete. To examine whether subjects overreact or under-react to wage raises, I compare subjects' effort levels in the last two periods in Treatment LH to those in Treatment HH, shown in part (a) of the figure. Similarly, I compare subjects' effort levels in the last two periods in Treatment HL to those in Treatment LL to examine subjects' responses to wage cuts, shown in part (b) of the figure. To control for the wealth effect, which may arise when subjects in treatment groups and control groups are paid differently in the first two periods, subjects will be paid by a randomly drawn period. All sessions are private to avoid the audience and coaction effect (Ariely et al., 2009).[4] To prevent subjects in different treatments from having different expectations about the direction of wage changes, subjects in all four treatments receives the following instruction: "*The piece wage rate will remain the same for the first two periods and may or may not change in Period 3. After that, it will remain the same in Period 4 as in Period 3.*"

2.2 Theoretical Predictions

[4] The presence of passive onlookers or spectators as well as others, who engage in the same activity, can significantly influence people's performance (Ariely et al., 2009).

Consider a standard model of effort provision: an individual's utility function is depicted as $v(x, e) = u(x) - c(e)$, where x is consumption bundle and e is effort. $c(e)$ is the effort function, where $c'(e) > 0$ and $c''(e) > 0$. Assume $u(x) = w \cdot e$ for simplicity.[5] w denotes the wage rate; $w \cdot e$ equals their earnings. The utility function can be rewritten as $v(x, e) = w \cdot e - c(e)$. Using the first-order necessary condition, the optimal effort level can be solved by $w = c'(e)$. The optimal effort depends solely on the wage rate, implying that individuals should neither overreact nor under-react to wage changes. For the rest of this section, I develop two behavioral hypotheses pointing in the opposite directions: one predicts subjects' under-reaction to wage change and the other predicts overreaction. The experiment therefore serves as a test for these two competing hypotheses.

First, recent studies in the application of prospect theory (Tversky and Kahneman, 1991) on labor supply behavior suggest that individuals' effort provision may exhibit reference-dependent preferences (for example, Camerer et al., 1997). I derive my hypothesis using the models of Koszegi and Rabin (2006) and Abeler et al. (2011), in which an individual derives "consumption utility" from consumption bundle x, and "gain-loss utility" from comparing her bundle to a reference bundle r. Overall utility is the sum of consumption and gain-loss utility. In addition, following Crawford and Meng (2010), I consider three cases in which individuals have income targets, effort targets, and targets for both. I find that, in any of the three cases, individuals would weakly under-react to wage changes. Figure 2 illustrates the theoretical predictions. Formal proof is left in Appendix B.

Hypothesis 1 (Under-reaction): Average effort in the last two periods of the LH treatment is not higher than that of the HH treatment. Similarly, average effort in the last two period of the HL treatment is not lower than that of the LL treatment.

The second behavioral hypothesis stems from a stream of psychological literature on the successive contrast effect. Extensive studies in this area have been done to explore the behavioral responses to changes in reward magnitude. Although the subjects are

---

[5] This assumption ensures positively-sloped labor supply curve. This assumption is plausible when outcomes in the setup are not very large (Abeler et al., 2011).

mainly animals, there is substantial evidence suggesting that behavior changes abruptly when reward quantity or quality is shifted, and that prior experience with rewards of different quantities and qualities affects the incentive value of a current reward. A pioneering study by Crespi (1942) showed that rats for whom reward was abruptly changed — either increased or decreased — altered their behavior rapidly; this behavior did not simply adjust to levels appropriate for the new rewards, but instead went beyond these levels (Flaherty, 1999). One of the main interpretations is based on emotional responses (Tinklepaugh, 1928; Spence, 1937; Crespi, 1942; Bower, 1961). Incentive relativity triggers elation effect or depression effect and thus results in overreaction. In the context of my experiment, subjects' prior exposure to a higher or lower wage rate influences how they perceive their current wage rates. In Period 3, subjects in the LH treatment may perceive the high wage rate more favorably than those in the HH treatment and may be more willing to work hard. Conversely, subjects in the HL treatment may perceive the low wage rate less favorably than those in the LL treatment and feel disappointed and thus unwilling to work.

Hypothesis 2 (Over-reaction): Average effort in the last two period of the LH treatment is higher than that of the HH treatment. Similarly, Average effort in the last two period of the HL treatment is lower than that of the LL treatment.

III. Results

A total of 123 undergraduate students at Brown University participated in the experiment during the period from September, 2012 to March, 2013. No subject participated in the experiment more than once. All sessions were private and lasted for roughly 20 minutes. Each session proceeded as follows: once subjects signed the consent form, they were brought into the computer lab. After they read the computerized instructions and completed the 15-second practice period, the main experiment began (please refer to Appendix A for experimental instructions). There was a 15-second break between periods. A post-experiment questionnaire is given to collect information on demographic characteristics and the rationale behind their decisions. Payments were made in cash before they left. The average monetary earnings were USD$10.10, including the guaranteed $5 participation fee.

Table 1 summarizes the descriptive statistics of the demographic variables. 30 or 31 subjects participated in each treatment. There is no significant difference between the values of any of the demographic characteristics for the four treatments, so a random assignment is valid *ex post*. Of the subjects, 48% were male. Half of the subjects categorized themselves as Caucasian, a third of the subjects as Asian, and the rest is Black, Native American or other. 27.3% of the subjects were economics-related concentrators. As subjects only participated in one treatment, the empirical results relied on between-subject variation.

Result 1: Subjects' work effort levels are higher when they are paid more.

Table 2 summarizes average work effort across the four treatments. Work effort level is simply the number of XZ pairs completed in a period. The average work effort level is between 880.9-899.5 pairs when the subjects are paid at the high rate in the first two periods (both HH and HL treatments) and is between 749.2-794.9 pairs when they were paid at the low rate in the first two periods (both LH and LL treatments). The difference is statistically significant according to the p-values of two-sided Mann-Whitney tests, indicating that subjects work harder when they are paid at a high wage rate.[6] This result is in line with the findings in Ariely et al. (2009), in which subjects performed a very similar keyboard-pressing task. However, it should be noted that the difference in the average work effort level is merely 13.7% even though the high wage rate is 10 times the low wage rate. The standard deviation is between 148.9-278.9, implying that both ability and willingness to press (disutility from pressing) may vary among the subjects.

Result 2: In Treatment HH and LL, subjects' work effort levels are stable in the first three periods, but increase significantly in the last period.

Figure 3 illustrates the changes in work effort across the four periods. In Treatments LL and HH, work effort is relatively stable in the first three periods, suggesting that neither the learning nor the fatigue effect – which would lead to an upward or downward trend – is significant in the keyboard-pressing task. The last period

[6] The p-values of two-sided Mann-Whitney test between any two treatments are all less than 0.05, except the one for Period 2 of the HH treatment vs. LH treatment, which is 0.09.

effect has occurred: effort level increases significantly in the final period, implying that the subjects exerted more energy toward the end of the experiment.[7]

Result 3: Subjects overreact to wage increases, but under-react to wage decreases.

To investigate whether subjects over- or under-react to wage increases, I compare subjects' effort levels in the last two periods of the HH and LH treatments, during which all subjects are paid at the high wage rate. In Period 3, the average effort level of the HH treatment is 884.0 pairs, whereas that of the LH treatment is 958.1 pairs, a difference of 74.1 pairs. The observed effort levels are marginally significant different according to the two-sided Mann-Whitney test (p-value = 0.898). However, in Period 4, the difference in average effort level between these two treatments falls to 52.1 pairs (= 983.5-931.4) and the effort levels are not significantly different (p-value = 0.245). These results suggest that individuals over-react to wage increases, but this effect diminishes with time.

Similarly, I compare the effort levels in the last two periods of the LL and HL treatments to examine whether subjects over- or under-react to wage decreases. In Period 3, the average effort level of the LL treatment is 765.4 pairs, whereas that of the HL treatment is 863.1 pairs. In Period 4, the effort levels are 791.3 pairs and 873.5 pairs in the LL and HL treatments, respectively. The difference is marginally significant for these two periods (p-value=0.086 in Period 3 and 0.075 in Period 4), indicating that subjects under-react to the wage decrease.

Bracha and Gneezy (2012) found that subjects over-reacted to changes in the show-up fee, both when it increased and decreased, whereas I find evidence to the contrary. Subjects in my experiment also overreact to wage increases; surprisingly, however, they under-react to wage decreases. After subjects are exposed to the high wage rate, the income target seems to influence their effort provision throughout the experiment. As shown in Figure 4, the individual-level data (in contrast to the aggregate level data) indicates that half of the subjects in the HL treatment increase their effort level in Period 3 after experiencing wage decrease. Even those who decrease their effort level mostly do so rather mildly (compared to the average effort level in the LL treatment). In

[7] Average effort level in Period 4 is higher than that in Period 3 by 27.4 pairs. The coefficient of Period 4 dummy in the regression, which will be discussed later, is significant at the 1% level.

addition to the quantitative analysis, their responses in the post-experiment questionnaire show some qualitative evidence suggesting an explanation for this:

Subject A: *"The Less Money per pair offered the faster you have to go so the more I concentrated."*

Subject B: *"When there was more money involved, I tried to get as many pairs as possible. When there was less money, I tried harder because I felt like it was a harder goal to accomplish and I wanted to see how much I could make, despite knowing I would make less money."*

Indeed, there are still some subjects who over-reacted to wage decreases. For example, a subject answered: *"I decided that it was not worth my time to engage meaningfully at a wage that would not conceivably put me at over a dollar\'s worth of earnings."* However, this type of subject only accounts for a small proportion among all the subjects in the HL treatment.[8]

The combination of the above results suggests that: first, consistent with Bracha and Gneezy's (2012) findings, past wages may act as a reference point and affect subjects' effort provision decisions. Standard economic models seem to overlook the behavioral factors which could be triggered in the face of changes. Second, unlike Bracha and Gneezy's (2012) findings, my subjects' responses to wage changes are not symmetric. They overreact to wage increases but under-react to wage decreases. The direction of change, whether up or down, matters significantly.

Result 4: Male subjects' average effort is higher than that of female subjects. However, the wage increase and the wage decrease both have a more pronounced effect on female subjects than their male counterparts.

I use regression analysis to explore how subjects respond to wage changes. Since the regressors are not correlated with unobserved variables (mainly the ability of the subject), the random effects model is more efficient than the fixed effects model. Regression results are shown in Table 3. The dependent variable, *effort level*, is simply

[8] Only 4 out of 30 subjects in the HL treatment reduce their effort levels by more than 120 pairs in Period 3. These responses may be loosely classified as over-reaction as the average effort level in the HH treatment is roughly 120 pairs higher than that in the LL treatment.

the number of XZ pairs completed in a period. *Wage* is an indicator taking value 1 if a subject in the period is paid at the high wage rate, and otherwise equaling 0. *Increase* is the dummy indicating whether a subject experiences wage increases. It takes the value of 1 for the observations in the last two periods of Treatment LH and 0 otherwise. Similarly, *decrease* is the dummy indicating whether a subject experiences wage decreases, which takes the value of 1 for the observations in the last two periods of Treatment HL and 0 otherwise. *Period2 - Period4* are period dummies to control for possible time effects. The rest are demographic variables.

In Column 1, I add W*age*, *Increase* and *Decrease* in the model. The coefficient of *Wage* is positively significant at the 1% level. This is consistent with Ariely et al.'s (2009) results and indicates that subjects work harder by pressing the XZ buttons more quickly when they are paid at the high wage rate. Although only the coefficient of Decrease is significant, the large positive estimates of *Increase* and *Decrease* suggest that, consistent with the results mentioned previously, subjects overreact to wage increases but under-react to wage reductions. The increasing but small positive estimates of *period2* and *Period3* suggests a slight upward trend in the effort level. *Period4* is significantly positive at the 1% level, indicating the last period effect.

I then add the demographic variables to the model, including Gender (*Male*), Race (*Asian* and *Caucasian*), Ethnicity (*Hispanic*), Concentration (*Econ*), and the number of semesters in college (*Semester*) in Column 2.[9] Only *Male* is statistically significant. The positive coefficient indicates that male subjects press the keyboard much quicker than their female counterparts. In Column 3, I investigate whether there is a gender difference in the responses to wage changes by adding the interaction terms: *Male*Increase* and *Male*Decrease*. Interestingly, although only *Male*Decrease* is marginally significant, the coefficients are both negative. This result shows that the wage increase and the wage decrease both have a more pronounced effect on female subjects than their male counterparts, suggesting that female subjects exhibit mildly stronger behavioral biases in

[9] *Male*, *Asian*, *Caucasian*, *Hispanic*, *Econ* are the dummy variables which indicates whether a subject is male, Asian Caucasian, Hispanic, and an economics-related major, respectively. *Semester* indicates the number of semesters a subject has been in college.

the face of wage changes.[10] The coefficients of *Increase* and *Decrease* become significant at the 5% level after controlling for demographic variables and interaction effects.

VI. Discussion and Conclusion

This paper experimentally examines whether people over- and under-react to wage changes. I design a simple, real-effort experiment in which subjects engage in a tedious keyboard-pressing task. Based on between-subject variation among the four treatment groups, the results suggest that subjects overreact to wage increases: subjects who experienced wage increases work harder than those who had been paid at the high wage rate. In contrast, subjects under-react to wage decreases: subjects who experienced wage decreases work harder than those who had been paid at the low wage rate. In addition, I also find that, interestingly, the wage increase and the wage decrease both have a more pronounced effect on female subjects than their male counterparts.

Before I extrapolate implications from my results, several questions need to be addressed. First, this research shares the same limitation as most laboratory experiments — lack of external validity (Kessler and Vesterlund, 2012). Even though it is a real-effort experiment, it still lacks some elements of the real-world working environment. For example, the concept of "leisure in the laboratory" (just pressing the keyboard slowly or doing nothing) is absent in the real world. The effort levels measured in this research are based on the speed with which the subjects press the keyboard but cannot show the amount of time they want to devote to working. Moreover, the duration of the experiment is so short that over- or under-reaction responses are statistically significant. However, it cannot be inferred from my data how persistent these irrational responses would be in the real world.

The results have two essential implications for research in labor supply. Firstly, the results suggest that some behavioral factors, such as emotions or cognitive biases, which significantly affect people's work effort decisions in the face of wage changes,

[10] I also add the interaction term between gender and wage rate (*Male*Wage*) to investigate the gender difference in their responses to wage rate. However, the coefficient is not significant in any specifications.

may have been overlooked in standard economic models. It seems to be inadequate to analyze people's responses to wage changes solely based on the substitution effect and the income effect. Ignoring these important factors may lead to inaccurate estimates of the sensitivity of labor supply to wage rates on which many fiscal policy makers may rely when evaluating tax and transfer programs.

Secondly, the possibility that subjects' responses to wage changes are direction-dependent has implications for existing studies on the wage elasticity of labor supply. If labor supply elasticity estimates are based on labor supply responses to incentive increases, such as a drop in marginal tax rate, this estimate cannot be applied to situations in which the incentive decreases. There is an extensive literature focusing on measuring wage elasticity of labor supply, some of which is based on natural experiments involving exogenous changes in the after-tax wage rate. My result implies that the estimates from wage-decreasing situations cannot be extrapolated to wage-increasing situations. The result also contributes to explain the highly variant estimates of wage elasticity of labor supply (Borjas, 2005; Evers et al., 2008).

The finding that people overreact to wage increases provides supporting evidence for the deferred compensation hypothesis which suggests an upward sloping age-earnings profile: workers are paid less than their marginal product of labor in the beginning of their career, with higher wages if they retain their jobs (Lazear, 1981). As people overreact to wage increases, the increasing wage profile may be an effective way to induce more effort. This finding may shed some light on why the deferred compensation approach is more commonly used by employers than the efficiency wage approach (Macpherson, Prasad and Salmon, 2011).

My finding that subjects under-react to wage decreases seems to contradict prior studies on wage rigidity, which suggests that cutting wages may have a large negative effect on work morale and thus companies tend to layoff workers rather than cut employees' wages during recessions (Bewley, 1999). If my results are accurate, the question is why? I speculate that since subjects were not allowed to leave the room before the experiment ends and they had nothing else to do in the laboratory, they seek some goals they can target. Thus, a substantial number of the subjects decided to work hard. Also, most workers in the real world have co-workers, and the way in which their co-workers experience and respond to wage changes might impact an individuals' response. These ideas are preliminary and still demand more research.

Reference

Abeler, J., Falk, A., Goette. L., and Huffman D. (2011), " Reference Points and Effort Provision." *American Economic Review*, 101(2): 470 –92.

Ariely, D., Bracha, A. and Meier, S. (2009), "Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially," *American Economic Review*, 99(1): 544-555.

Ariely, D., Gneezy, U., Lowenstein G., and Mazar N.(2009), "Large Stakes and Big Mistakes," *Review of Economic Studies*, 76, 451–469

Bewley, T. (1999), *Why Wages Don't Fall During a Recession*, Harvard University Press.

Blundell, R. and MaCurdy, T., (1999) "Labor Supply: A Review of Alternative Approaches," edited by O.C. Ashenfelter and D. Card, *Handbook of Labor Economics*, Volume 3A, Amterdam: Elsevier, 1559-1695.

Borjas, J. G., (2012), *Labor Economics*, McGraw-Hill/Irwin.

Bower, G. H. (1961), "A Contrast Effect in Differential Conditioning," *Journal of Experimental Psychology*, 62, 196-199.

Bracha, A. and Gneezy, U. (2012) "Relative Pay and Labor Supply," Working Paper http://www.bos.frb.org/economic/wp/wp2012/wp1206.pdf

Camerer, C., Babcock, L., Loewenstein, G. and Thaler, R., (1997), "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics*, 112(2): 407-441.

Crawford, V. and Meng, J. (2011), "New York City Cabdrivers.Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *American Economic Review*, 101, 1912-1932.

Crespi, L. P. (1942), "Quantitative variation in incentive and performance in the white rat," *American Journal of Psychology*, 40: 467–517.

Ericson, K. and Fuster, A. (2011), "Expectations as Endowments: Evidence on Reference-Dependent Preferences from Exchange and Valuation Experiments," *Quarterly Journal of Economics*, Vol. 126, 4, 1869-1907.

Evers, M., De Mooij, R. and Van Vuuren, D. (2008), "The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates," *De Economist*, 156: 25-43.

Fehr, E. and Goette, L. (2007), "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, 97(1):298-317.

Feldstein, M., and Poterba, J. (1984), “Unemployment Insurance and Reservation Wages,” *Journal of Public Economics,* 23(2): 141–167.

Flaherty, C. (1996),”Brief History of Reward Magnitude Research,” *Incentive Relativity*, Cambridge University Press, Cambridge, U.K., p. 5-18.

Heath, C., Larrick, R. and Wu, G. (1999), “Goals as Reference Points,” *Cognitive Psychology,* 38, 79–109.

Kessler J. and Vesterlund L. (2012), "The External Validity of Laboratory Experiments: Qualitative rather than Quantitative Effects," in “*Methods of Modern Experimental Economics*”, edited by Guillaume Frechette and Andrew Schotter, Oxford University Press.

Koszegi, B. and Rabin, M. (2006), “A Model of Reference-Dependent preferences,” *Quarterly Journal of Economics*, 121(4), 1133-1165.

Lazear, E. (1981), “Agency, earnings profiles productivity and hours restrictions,” *American Economic Review,* 71: 606-620.

Macpherson, D., Prasad, K. and Salmon, T.C., (2011) “Deferred Compensation Vs. Efficiency Wages: An Experimental Test of Effort Provision and Self-Selection,” Working Paper. http://faculty.smu.edu/tsalmon/MPS.pdf

McClelland, R. and Mok, S. (2012) “A Review of Recent Research on Labor Supply Elasticities,” Working Paper Series, Congressional Budget Office, Washington, D.C. http://www.cbo.gov/sites/default/files/cbofiles/attachments/10-25-2012-Recent_Research_on_Labor_Supply_Elasticities.pdf

Spence, K.W. (1937), The Differential Response of Animals to Stimuli Differing within a Single Dimension, *Psychological Review*, 44, 430-444.

Tinklepaugh, O.L. (1928), “An Experimental Study of Representative Factors in Monkeys,” *Journal of Comparative Psychology*, 8, 197-236.

Tversky, A. and Kahneman D., (1991) “Loss Aversion in Riskless Choice: A Reference-Dependent Model,” *The Quarterly Journal of Economics*, 106(4): 1039-1061.

Tables and Figures

Table 1: The summary of demographic variables

Treatment	# of Obs	Gender	Asian	Caucasian	Hispanic	Econ	Semesters
HH	31	48.4%	32.3%	58.1%	19.4%	20.0%	2.8
HL	30	50.0%	36.7%	46.7%	20.0%	31.0%	2.5
LH	31	48.4%	35.5%	41.9%	25.8%	22.6%	2.5
LL	31	45.2%	29.0%	48.4%	19.4%	35.5%	2.6
All	123	48.0%	33.3%	48.8%	21.1%	27.3%	2.6

Note: One subject in the HH treatment had unreported *Econ* and *Semester*. One subject in the HL treatment had unreported Econ. Mann-Whitney tests are used to test the difference of means. None of the means is significantly different at the 10% level.

Table 2: Work efforts in four treatments

Treatment	Period 1	Period 2	Period 3	Period 4
HH	899.5	890.5	884.0	931.4
	(182.8)	(186.3)	(187.7)	(192.9)
HL	898.2	880.9	863.1	873.6
	(221.9)	(236.4)	(244.2)	(278.9)
LH	771.0	794.9	958.1	983.5
	(152.8)	(161.7)	(181.8)	(202.0)
LL	749.2	765.3	765.4	791.3
	(148.9)	(169.5)	(182.5)	(190.7)
All	828.9	832.5	867.7	895.1
	(189.8)	(195.6)	(209.7)	(227.6)

Notes: Standard errors reported in parentheses. Work effort is measured by the number of XZ pairs completed in a period.

Table 3: Regression

	(1)	(2)	(3)
Wage	120.39***	106.72***	106.93***
	(33.68)	(32.08)	(31.94)
Increase	52.11	64.72*	80.30**
	(36.44)	(35.00)	(37.36)
Decrease	82.67**	67.22*	91.50**
	(36.53)	(35.11)	(37.67)
Male		96.56***	106.65***
		(32.18)	(32.42)
Male*Increase			-32.70
			(28.15)
Male*Decrease			-49.99*
			(29.06)
Period2	3.59	4.47	4.47
	(10.12)	(10.27)	(10.25)
Period3	4.49	6.22	6.25
	(12.27)	(12.44)	(12.41)
Period4	31.95***	33.23***	33.25***
	(12.27)	(12.44)	(12.41)
Asian		+	+
Caucasian		+	+
Hispanic		-	-
Econ		+[a]	+[a]
Semester		+[a]	+[a]
Constant	775.20***	688.50***	668.48***
# of Observations	492	484	484
R-square	0.34	0.34	0.34

Notes: Random effects OLS regressions. Standard errors reported in parentheses. Dependent variable is the number of XZ pairs completed in a period.

*** indicates significance at the 1% level;
** indicates significance at the 5% level;
* indicates significance at the 10% level.

[a] Two subjects had unreported demographic characteristics.

Figure 1: The four treatment groups

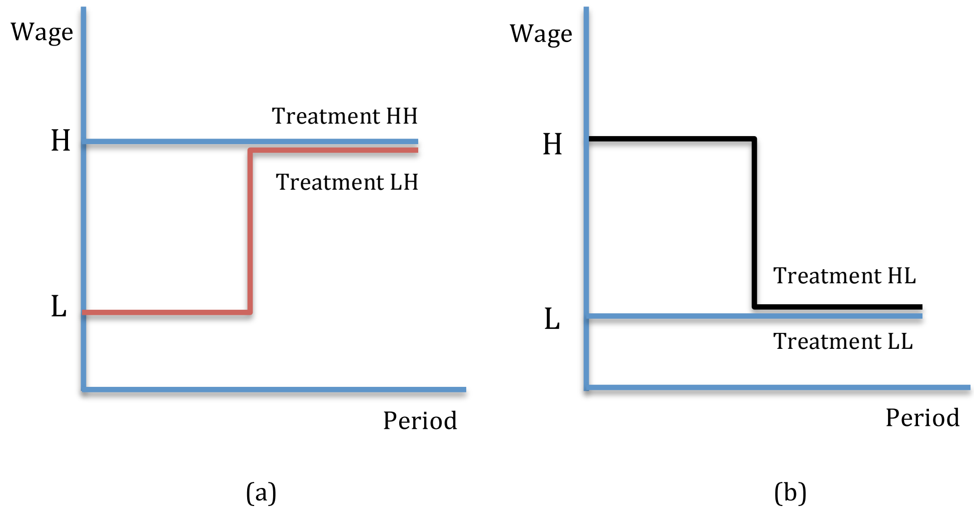


Figure 2: Theoretical Predictions

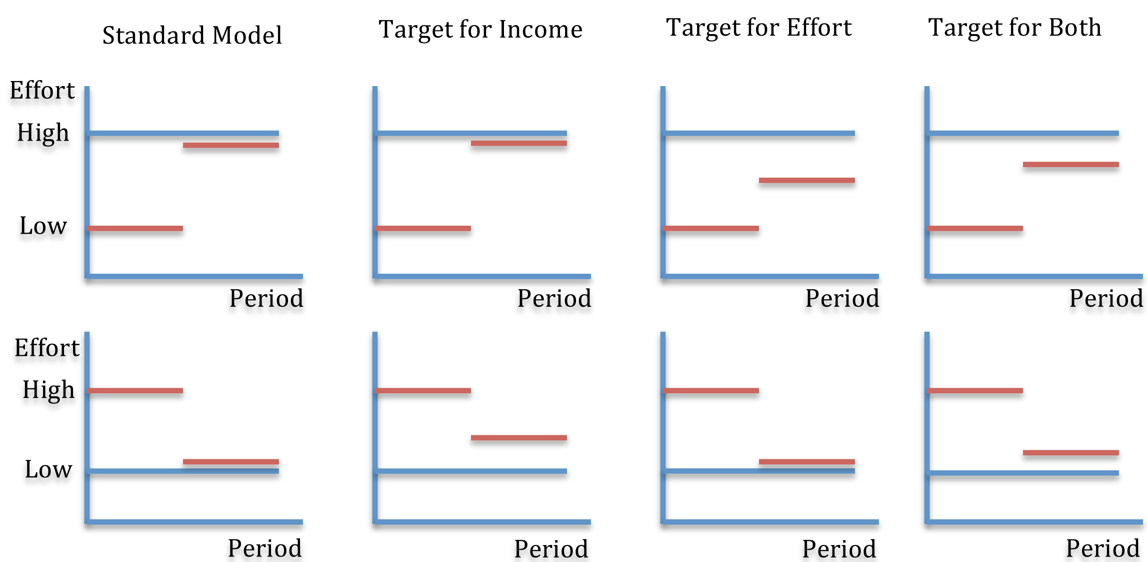
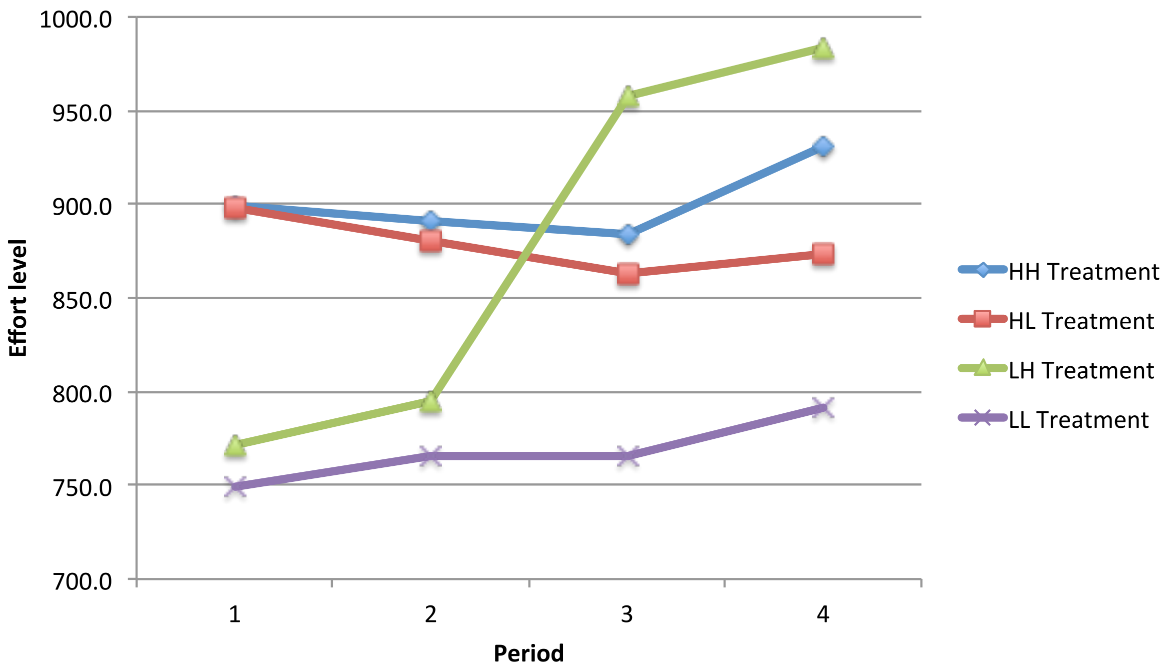
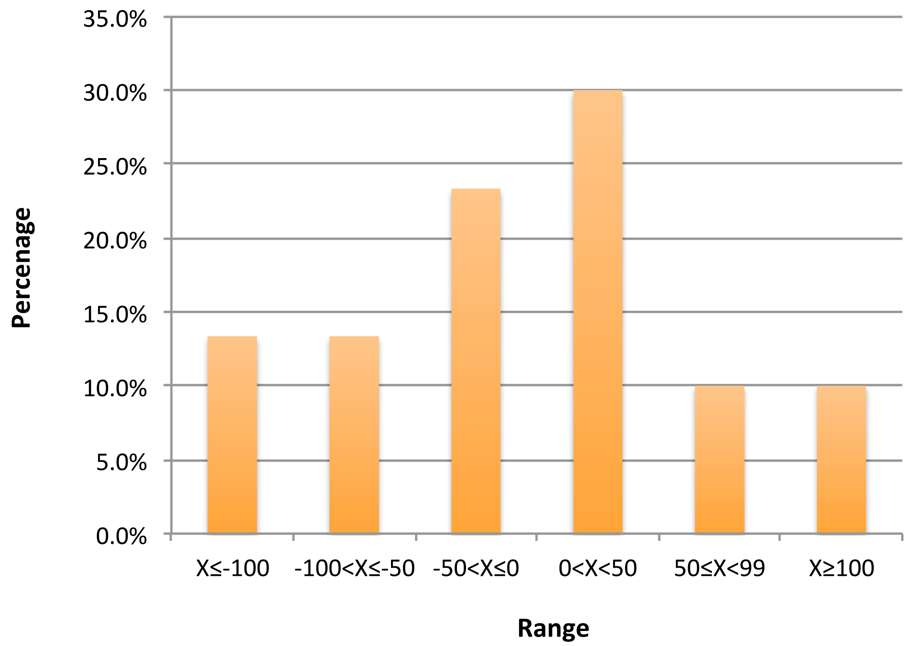


Figure 3: Effort levels in the four treatments



Note: Effort level is measured by the number of pairs a subject completed in a period.

Figure 4: Subjects' responses to wage decreases



Notes:
1. X is the difference in a subject's effort levels between Period 2 and Period 3. If X is positive, the subject increases her effort level in response to the wage increase. Otherwise, the subject decreases her effort level in response to the wage increase.
2. 30 subjects in the HL treatment.

Appendix A: Instructions

In the beginning of the experiment:

Thanks for participating in this experiment. The experiment will last for 20 minutes. NO communication between participants is allowed at any time during the experiment. If you have any questions, please raise your hand and I will come to assist you. Please now turn off your mobile phone and any other electronic devises. These must remain off until you leave this room.

The experiment consists of four periods in addition to a shorter 15-second practice non-paying period which aims to formalize you with the task. In each period you will undertake an identical task lasting 4 minutes. Your task is to sequentially press “X” and “Z” on the keyboard. When performing the task, both in the practice period and the actual four periods, you should use one hand only.

Your period earnings equal the number of X-Z pairs you complete times the piece wage rate in the period. For example, if you complete 1000 pairs and the piece wage rate is X cents per pair. Then, your earnings in this period are 1000X cents = 10X dollars.

You will see the applicable piece wage rate of a period 5 seconds before the period starts. The piece wage rate will remain the same for the first two periods and may or may not change in Period 3. Regardless of whether the piece rate wage changes in Period 3, it will be the same in Period 4 as in Period 3.

At the end of the experiment, you will be paid by a randomly chosen period plus the $5 guaranteed for participation. Because there are four periods in total, the probability that a period will be selected is 1/4 or 25%.

In the beginning of Period 1:

In this period, you will be paid $_____ per pair. The rate will remain the same for the first two periods.

In the beginning of Period 2:

In this period, you will be paid $_____ per pair.

In the beginning of Period 3:

In this period, you will be paid $_____ per pair. The rate will remain the same for the remaining two periods.

In the beginning of Period 4:

In this period, you will be paid $_____ per pair.

Appendix B: Theoretical Model

In this appendix, I follow Crawford and Meng's (2010) setting and consider three different targets an individual may have. Suppose an individual exhibits reference-dependent preferences and s/he may be paid at the high wage rate w_H or the low wage rate w_L. Define e_H^* and e_L^* as the optimal effort level which can be solved by the first-order necessary conditions, $w_H = c'(e)$ and $w_L = c'(e)$, respectively.

Case 1: Target on Income

An individual maximizes the reference-dependent utility function:

$$Max\ \ v(x,e) = (1-\eta)u(x) + \eta[u(x) - u(x^r)] \cdot \lambda - c(e)$$

$$\lambda > 1, if x < x^r$$

$$\lambda = 1, if x \geq x^r$$

e is the effort level; x^r is the income target; η is the parameter which represents the relative magnitude of the loss-gain utility as opposed to the consumption utility, $0 \leq \eta < 1$; and λ represents how much extra cost an individual incurs if the income target is not achieved, $\lambda > 1$. Assume $u(x) = x = w \cdot e$, where *w* is the wage rate, we can rewrite the utility function as:

$$v(x,e) = (1-\eta)w \cdot e + \eta[w \cdot e - u(x^r)] \cdot \lambda - c(e)$$

The Kuhn-Tucker first-order necessary condition is:

$$(1-\eta+\eta \cdot \lambda) \cdot w = c'(e), if\ x < x^r \tag{B.1}$$

$$w = c'(e), if\ x \geq x^r \tag{B.2}$$

Note that $(1-\eta+\eta \cdot \lambda) \geq 1$. The equation (B.1) implies that, for a given wage rate *w*, the optimal level of effort is higher than that predicted by the standard model ($w = c'(e)$) if the income target is not achieved. The intuition behind this result is that since individuals would incur extra cost if s/he does not achieve the income target, s/he would exert more effort.

In my experiment, a subject may be paid at the high wage rate w_H or the low wage rate w_L. Consider the subjects in the HL treatment. In the first two periods when the subjects are paid at the high wage rate, they form a high income target, $x^r = w_H \cdot e_H^*$. In Period 3, since the wage rate is decreased to w_L , the Kuhn-Tucker first-order condition becomes:

$$(1-\eta+\eta\cdot\lambda)\cdot w_L = c'(e), if\ w_L\cdot e < w_H\cdot e_H^* \tag{B.3}$$

$$w_L = c'(e), if\ w_L\cdot e \geq w_H\cdot e_H^* \tag{B.4}$$

Equation (B.4) is contradictory because $w_L \cdot e > w_H \cdot e_H^*$. The optimal effort level (e_L^{**}) is the effort level which solves equation (B.3). Since $(1-\eta+\eta\cdot\lambda) > 1$, $e_L^{**} > e_L^*$, implying subjects under-react to wage decreases.

It is trivial to show that the optimal levels in Period 3 of the other three treatments are the same as the standard theoretical predictions because the income target is achieved when the wage rate increases or remains the same.

Case 2: Target on Effort

An individual maximizes the reference-dependent utility function:

$$Max\ \ v(x,e) = u(x) - \{(1-\eta)c(e) + \eta[c(e) - c(e^r)]\cdot\lambda\}$$

$$\lambda > 1, if e > e^r$$

$$\lambda = 1, if e \leq e^r$$

where e^r is the effort target and λ represents how much extra cost an individual incurs if the effort target is higher than the effort target, $\lambda > 1$. Assume $u(x) = w\cdot e$, we can rewrite the utility function as:

$$v(x,e) = w\cdot e - \{(1-\eta)c(e) + \eta[c(e) - c(e^r)]\cdot\lambda\}$$

The Kuhn-Tucker first-order condition is:

$$w = (1-\eta+\eta\cdot\lambda)\cdot c'(e), if\ e > e^r \tag{B.5}$$

$$w = c'(e), if\ e \leq e^r \tag{B.6}$$

Note that $(1-\eta+\eta\cdot\lambda) \geq 1$. Equation (B.5) implies that, for a given wage rate *w*, the optimal level of effort is lower than that predicted by the standard model ($w = c'(e)$) if the individual exerts more effort than the target level. The intuition behind this result is that individuals would incur extra cost if s/he exerts more effort than the effort target.

In my experiment, a subject may be paid at the high wage rate w_H or the low wage rate w_L. Consider the subjects in the LH treatment. In the first two periods when the subjects are paid at the low wage rate, they form a low effort target, $e^r = e_L^*$. In Period 3, since the wage rate is increased to w_H, the Kuhn-Tucker first-order condition becomes:

$$(1-\eta+\eta\cdot\lambda)\cdot w_H = c'(e), if\ e > e_L^*$$

$$w_H = c'(e), if\ e \leq e_L^*$$

Because $w_H = c'(e)$ contradicts with $e \leq e_L^*$, the optimal effort level, e_H^{**}, is solved by $(1-\eta+\eta\cdot\lambda)\cdot w_H = c'(e)$. Since $(1-\eta+\eta\cdot\lambda) > 1$, $e_H^{**} > e_H^*$, implying subjects under-react to wage increases.

It is trivial to show that the optimal levels in Period 3 of the other three treatments are the same as the standard theoretical predictions because the effort target is not exceeded when the wage rate decreases or remains the same.

Case 3: Target on Both Income and Effort

An individual maximizes the reference-dependent utility function:

$$Max\ v(x,e) = (1-\eta)[u(x)-c(e)] + \eta[u(x)-u(x^r)]\cdot\lambda_1 + \eta[c(e^r)-c(e)]\cdot\lambda_2$$

$$\lambda_1 > 1, if x < x^r$$

$$\lambda_1 = 1, if x \geq x^r$$

$$\lambda_2 > 1, if e > e^r$$

$$\lambda_2 = 1, if e \leq e^r$$

$$\lambda_1 = \lambda_2$$

where x^r is the income target and e^r is the effort target. $\lambda_1, \lambda_2 > 1$ Assume $u(x) = w\cdot e$, we can rewrite the utility function as:

$$Max\ v(x,e) = (1-\eta)[w\cdot e-c(e)] + \eta[w\cdot e-u(x^r)]\cdot\lambda_1 + \eta[c(e^r)-c(e)]\cdot\lambda_2$$

The Kuhn-Tucker first-order necessary condition is:

$$(1-\eta+\eta\cdot\lambda_1)\cdot w = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e), if\ x < x^r\ and\ e > e^r$$

$$(1-\eta+\eta\cdot\lambda_1)\cdot w = c'(e), if\ x < x^r\ and\ e \leq e^r$$

$$w = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e), if\ x \geq x^r\ and\ e > e^r$$

$$w = c'(e), if\ x \geq x^r\ and\ e \leq e^r$$

Consider the subjects in the LH treatment. In the first two periods when the subjects are paid at the low wage rate, they form a low income target, $x^r = w_L \cdot e_L^*$, and a low effort target, $e^r = e_L^*$. In Period 3, since the wage rate is increased to w_H, the first-order condition becomes:

$$(1-\eta+\eta\cdot\lambda_1)\cdot w_H = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e), if\ w_H\cdot e < w_L\cdot e_L^*\ \ and\ e > e_L^* \quad \text{(B.7)}$$

$$(1-\eta+\eta\cdot\lambda_1)\cdot w_H = c'(e), if\ w_H\cdot e < w_L\cdot e_L^*\ and\ e \leq e_L^* \quad \text{(B.8)}$$

$$w_H = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e), if\ w_H\cdot e \geq w_L\cdot e_L^*\ and\ e > e_L^* \quad \text{(B.9)}$$

$$w_H = c'(e), if\ w_H\cdot e \geq w_L\cdot e_L^*\ and\ e \leq e_L^* \quad \text{(B.10)}$$

It can be shown that only equation (B.9) is consistent. Therefore, e_H^{**} solves $w_H = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e)$. So, $e_H^{**} < e_H^*$, implying subjects under-react to wage increases.

Consider the subjects in the HL treatment. In the first two periods when the subjects are paid at the high wage rate, they form a high income target, $x^r = w_H \cdot e_H^*$, and a high effort target, $e^r = e_H^*$. In Period 3, since the wage rate is decreased to w_L, the first-order condition becomes:

$$(1-\eta+\eta\cdot\lambda_1)\cdot w_L = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e), if\ w_L\cdot e < w_H\cdot e_H^*\ \ and\ e > e_H^* \quad \text{(B.11)}$$

$$(1-\eta+\eta\cdot\lambda_1)\cdot w_L = c'(e), if\ w_L\cdot e < w_H\cdot e_H^*\ and\ e \leq e_H^* \quad \text{(B.12)}$$

$$w_L = (1-\eta+\eta\cdot\lambda_2)\cdot c'(e), if\ w_L\cdot e \geq w_H\cdot e_H^*\ and\ e > e_H^* \quad \text{(B.13)}$$

$$w_L = c'(e), if\ w_L\cdot e \geq w_H\cdot e_H^*\ and\ e \leq e_H^* \quad \text{(B.14)}$$

It can be shown that only equation (B.12) is consistent. Therefore, e_L^{**} solves $(1-\eta+\eta\cdot\lambda_1)\cdot w_L = c'(e)$. So, $e_L^{**} > e_L^*$, implying subjects also under-react to wage decreases.

It is trivial to show that the optimal levels in Period 3 of the other two treatments (HH and LL treatments) are the same as the standard theoretical predictions because neither the income target is achieved nor the effort target is exceeded when the wage rate remains the same.